On error distance of received words with fixed degrees to Reed-Solomon code[*]

Li Yujuan
Science and Technology on Information Assurance Laboratory
Beijing, P.R.China
liyj@amss.ac.cn

Zhu Guizhen
Data Communication Science and Technology Research Institute
Beijing, P. R.China
zhugz08@gmail.com

Abstract

Under polynomial time reduction, the maximum likelihood decoding of a linear code is equivalent to computing the error distance of a received word. It is known that the decoding complexity of standard Reed-Solomon codes at certain radius is at least as hard as the discrete logarithm problem over certain large finite fields. This implies that computing the error distance is hard for standard Reed-Solomon codes. Using some elegant algebraic constructions, we are able to determine the error distance of received words whose degree is $k+1$ to the Standard Reed-Solomon code or Primitive Reed-Solomon code exactly. Moreover, we can precisely determine the error distance of received words of degree $k+2$ to the Standard Reed-Solomon codes. As a corollary, we can simply get the results of Zhang-Fu-Liao and Wu-Hong on the deep hole problem of Reed-Solomon codes.

1 Introduction

Let $\mathbb{F}_q$ be the finite field with q elements, where q is a prime power. For positive integers $k < n \leq q$, the generalized Reed-Solomon code, denoted by $\mathcal{C}$, can be thought of as a map from $\mathbb{F}_q^k \longrightarrow \mathbb{F}_q^n$, in which a message $(a_0, a_1, \ldots, a_{k-1})$ is mapped to a vector $(f(x_1), f(x_2), \ldots, f(x_n))$, where $f(x) = a_{k-1}x^{k-1} + a_{k-2}x^{k-2} + \cdots + a_0 \in \mathbb{F}_q[x]$ and $D = \{x_1, x_2, \ldots, x_n\} \subseteq \mathbb{F}_q$ is called the evaluation set. It is obvious that $\mathcal{C}$ is a linear subspace of $\mathbb{F}_q^n$ with dimension k. When the evaluation set is the whole field $\mathbb{F}_q$, the resulting code is called the standard Reed-Solomon code, denoted by $\mathcal{C}_q$. If the evaluation set is $\mathbb{F}_q^*$, the resulting code is called primitive Reed-Solomon code, denoted by $\mathcal{C}_q^*$.

[*] A preliminary version can be seen in the first author's PhD thesis in 2008.

The Hamming distance between two codewords is the number of coordinates in which they differ. The error distance of a received word $u \in \mathbb{F}_q^n$ to the code $\mathcal{C}$ is the minimum Hamming distance of u to codewords, denoted by $d(u, \mathcal{C})$. A Hamming ball of radius m is the set of vectors within Hamming distance m to some vector in $\mathbb{F}_q^n$. The minimum distance of a code is the smallest distance between any two distinct codewords, and is a measure of how many errors the code can correct or detect. The covering radius of a code is the maximum possible distance from any vector in $\mathbb{F}_q^n$ to the closest codeword. A deep hole is a vector which achieves this maximum. The minimum distance of generalized Reed-Solomon codes is $n - k + 1$. The covering radius of generalized Reed-Solomon codes is $n-k$. Therefore, all the deep holes of Reed-Solomon code are the vectors of error distance $n - k$.

1.1 Related Work

The complexity for decoding Reed-Solomon codes has also attracted attention recently. Guruswami and Vardy [8] proved that the maximum likelihood decoding of generalized Reed-Solomon codes is NP-hard. In fact, the weaker problem of deciding deep holes for generalized Reed-Solomon codes is already co-NP-complete, see [4]. In the much more interesting case of standard Reed-Solomon codes, it is unknown if decoding remains NP-hard. This is still an open problem. Cheng and Wan [5] [6] managed to prove that the decoding problem of standard Reed-Solomon codes at certain radius is at least as hard as the discrete logarithm problem over a large extension of a finite field. This is the only complexity result that is known for decoding the standard Reed-Solomon code.

Under polynomial time reduction, the maximum likelihood decoding of a linear code is equivalent to computing the error distance of a received word. Our aim of this paper is to study the problem of computing the error distance of received words of certain degrees to the Reed-Solomon code. We shall use algebraic methods. For this purpose, we first define the notion of the degree of a received word. For $u = (u_1, u_2, \ldots, u_n) \in \mathbb{F}_q^n$, $D = \{x_1, \ldots, x_n\} \subset \mathbb{F}_q$, let

$$u(x) = \sum_{i=1}^{n} u_i \frac{\prod_{j \neq i}(x - x_j)}{\prod_{j \neq i}(x_i - x_j)} \in \mathbb{F}_q[x].$$

That is, $u(x)$ is the unique (Lagrange interpolation) polynomial of degree at most $n-1$ such that $u(x_i) = u_i$ for $1 \leq i \leq n$. For $u \in \mathbb{F}_q^n$, we define $\deg(u) = \deg(u(x))$, called the degree of u. It is clear that $d(u, \mathcal{C}) = 0$ iff $\deg(u) \leq k - 1$. Without loss of generality, we can assume that $k \leq \deg(u) \leq n - 1$ and $u(x)$ is monic. We have the following simple bound.

Lemma 1 *For $k \leq \deg(u) \leq n - 1$, we have the inequality*

$$n - \deg(u) \leq d(u, \mathcal{C}) \leq n - k.$$

This result shows that if $\deg(u) = k$, then $d(u, \mathcal{C}) = n - k$ and thus u is a deep hole. As mentioned before, it is NP-hard to determine whether $d(u, \mathcal{C}) = n - k$ (the deep hole problem) for generalized Reed-Solomon codes. Thus, one way to exploit this problem is restricting our attention to the most natural and important case, namely the standard Reed-Solomon code $\mathcal{C}_q$. Even in this restricted case, we cannot expect a complete solution to the problem of computing the error distance, as it is at least as hard as the discrete logarithm in a large finite field. However, we expect that a lot more can be said for standard Reed-Solomon codes. For instance, Cheng and Murray [4] conjectured the following complete classification of deep holes for standard Reed-Solomon codes.

Conjecture (Cheng-Murray). All deep holes for standard Reed-Solomon codes are those words satisfying $\deg(u) = k$. In other words, a received word u is a deep hole for $\mathcal{C}_q$ iff $\deg(u) = k$.

The deep hole problem for generalized Reed-Solomon codes is NP-hard. In contrast, the Cheng-Murray conjecture implies that the deep hole problem for the standard Reed- Solomon code can be solved in polynomial time. A complete proof of this conjecture (if correct) seems rather difficult at present. As a theoretical evidence, they proved that their conjecture is true if $d := \deg(u) - k$ is small and q is sufficiently large compared to $d + k$. More precisely, they showed

Proposition 1 *Let $u \in \mathbb{F}_q^q$ such that $1 \leq d := \deg(u) - k \leq q - 1 - k$. Assume that $q \geq \max\{k^{7+\epsilon}, d^{\frac{13}{3}+\epsilon}\}$ for some constant $\epsilon > 0$. Then $d(u, \mathcal{C}_q) < q - k$, that is, u is not a deep hole.*

However, they did not obtain the exact value of $d(u, \mathcal{C}_q)$, only the weaker inequality $d(u, \mathcal{C}_q) < q - k$. Li and Wan [10] improved their results using Weil's character sum estimate and the approach of Cheng-Wan [5] as follows.

Proposition 2 *Let $u \in \mathbb{F}_q^q$ such that $1 \leq d := \deg(u) - k \leq q - 1 - k$. For some constant $\epsilon > 0$,*

1) if

$$q \geq \max\{(k+1)^2, d^{2+\epsilon}\}, \text{ and } k > \left(\frac{2}{\epsilon} + 1\right) d + \frac{8}{\epsilon} + 2,$$

then u is not a deep hole.

2) if

$$q \geq \max\{(k+1)^2, (d-1)^{2+\epsilon}\}, \text{ and } k > \left(\frac{4}{\epsilon} + 1\right) d + \frac{4}{\epsilon} + 2,$$

then $d(u, \mathcal{C}_q) = q - (k + d)$.

Note that the last part of the proposition determines the exact error distance $d(u, \mathcal{C}_q)$ under a suitable hypothesis. Using a similar character sum approach, Qunying Liao [11] unified the above two results of Li-Wan and proved the following extension.

Proposition 3 *Let $r \geq 1$ be an integer. For any received word $u \in \mathbb{F}_q^q$,$r \leq d := \deg(u) - k \leq q - 1 - k$. If*

$$q \geq \max\left\{2\binom{k+r}{2} + d, d^{2+\epsilon}\right\}, \text{ and } k > \left(\frac{2}{\epsilon} + 1\right)d + \frac{4+2r}{\epsilon} + 2,$$

for some constant $\epsilon > 0$, then $d(u, \mathcal{C}_q) \leq q - k - r$.

Antonio Cafure etc. [2] uses a much more sophisticated algebraic geometry approach and obtains a slightly improvement of one of the Li-Wan results.

Proposition 4 *Let $u \in \mathbb{F}_q^q$ such that $1 \leq d := \deg(u) - k \leq q - 1 - k$. Assume that*

$$q \geq \max\{(k+1)^2, 14d^{2+\epsilon}\}, \text{ and, } k > \left(\frac{2}{\epsilon} + 1\right)d,$$

for some constant $\epsilon > 0$, then u is not a deep hole.

Again, this result gives only the inequality $d(u, \mathcal{C}_q) < q - k$, not the exact value of the error distance $d(u, \mathcal{C}_q)$. As for the error distance, Zhu-Wan [15] prove the following result.

Proposition 5 *Let $r \geq 1$ be an integer and $u \in \mathbb{F}_q^q$, $r \leq d := \deg(u) - k \leq q - 1 - k$. There are positive constants c_1 and c_2 such that if*

$$d < c_1 q^{1/2}, \left(\frac{d+r}{2} + 1\right)\log_2 q < k < c_2 q,$$

then $d(u, \mathcal{C}_q) \leq q - k - r$.

So far, no one has proved or defied Cheng-Murray's conjecture on standard Reed-Solomon code. In a recent paper by Cheng-Li-Zhuang[3], they classify deep holes completely for generalized Reed-Solomon codes $\mathcal{C}$, where q is prime, and $D > k \geq \frac{q-1}{2}$. Moreover, they prove that

Proposition 6 *Cheng-Murray's conjecture is true for $p > 2, k + 1 \leq p$ or $3 \leq q - p + 1 \leq k + 1 \leq q - 2$.*

Another way to research the maximum likely decoding problem or deep hole problem is studying the error distance of received words of certain degree to Reed-Solomon code. Wu-Hong[13] show that some received words of degree $q - 2$ are deep holes of Primitive Reed-Solomon code.

Proposition 7 *Let $\mathcal{C}_q^*$ be a Primitive Reed-Solomon code, $q \geq 4$, and $2 \leq k \leq q - 2$, then all the received words that can be presented by polynomial $u(x) = ax^{q-2} + v(x)$ are deep holes, where $\deg(v) \leq k - 1$, $a \neq 0$.*

Zhang-Fu-Liao[14] extend the result above to any evaluation set $D \neq \mathbb{F}_q$, and derive the following conclusion.

Proposition 8 *Let $\mathcal{C}$ be a Generalized Reed-Solomon code, $D \neq \mathbb{F}_q$, then for $a \neq 0$, $b \notin D$, all the received words that can be presented by polynomial $u(x) = a(x-b)^{q-2}+v(x)$ are deep holes, where $\deg(v) \leq k-1$.*

Meanwhile, they find another kind of deep holes for certain dimension and finite fields.

Proposition 9 *Let q be a power of 2, and $q \geq 4$, $\mathcal{C}$ be a Generalized Reed-Solomon code, evaluation set $D = \mathbb{F}_q^*$ or $D = \mathbb{F}_q^* \setminus \{1\}$, $k = q-4$. If $a \neq 0$, then all the received words that can be presented by polynomial $u(x) = ax^{q-3} + v(x)$ are deep holes, where $\deg(v) \leq k-1$.*

Finally, they prove that if $q > 5$ with odd characteristic p, and $2 \leq k \leq q-3$, all the received words represented by the following polynomials are not deep holes of Primitive Reed-Solomon code.

$$u(x) = ax^{k+2} + bx^{k+1} + cx^k + v(x),$$

where $a \in \mathbb{F}_q^*$, $b, c \in \mathbb{F}_q$, and $\deg(v) \leq k-1$.

Again, this result gives only the inequality $d(u, \mathcal{C}_q) < q-k$, not the exact value of the error distance $d(u, \mathcal{C}_q)$. And they only discuss the case that characteristic $p \neq 2$. For $p = 2$, the problem may be more complicated, which can be seen in our analysis in this paper.

1.2 Our results

In this paper, we focus on computing the error distance of received words of fixed degrees to Reed-Solomon codes. The main results consist of two parts. Firstly, we exploit the error distance of received words of degree $k+1$ to Standard Reed-Solomon code and Generalized Reed-Solomon code; Secondly, we compute the error distance of received words of degree $k+2$ to Standard Reed-Solomon code not only for $p \neq 2$ but also for the case $p = 2$. As a corollary, we can rather easily get Wu-Hong and Zhang-Fu-Liao's results on deep hole.

Theorem 1 *(i) Let $\mathcal{C}_q$ be a Standard Reed-Solomon code, and $u \in \mathbb{F}_q^n$ represented by polynomial $u(x) = x^{k+1} - bx^k + v(x)$, $\deg(v) \leq k-1$, then $d(u, \mathcal{C}_q) = q-k$ if one of the following holds*

(a) $b = 0, p = 2, k = 1$,

(b) $b = 0, p = 2, k = q-3$,

otherwise, $d(u, \mathcal{C}_q) = q-k-1$.

(ii) If $D = \mathbb{F}_q^*$ and $q > 5$, then $d(u, \mathcal{C}) = q-k-1$ if one of the following holds

(a) $b = 0, p = 2, k = 1$,

(b) $b = 0, p = 2, k = q-4$,

(c) $b = 0, k = q - 3$.

otherwise, $d(u, \mathcal{C}) = q - k - 2$.

Theorem 2 *Let $\mathcal{C}_q$ be a Standard Reed-Solomon code, $k \geq 1, k+2 \leq q-1$, and $u \in \mathbb{F}_q^n$ represented by polynomial $u(x) = x^{k+2} - bx^{k+1} + cx^k + v(x)$, $\deg(v) \leq k-1$, then*

(i) If $k + 2 = q - 1$, then

$$d(u, \mathcal{C}_q) = \begin{cases} q - k - 2 & \textit{if } b^2 = c, \\ q - k - 1 & \textit{if } b^2 \neq c. \end{cases}$$

(ii) If $p = 2$ and $k + 2 \leq q - 2$, then we can get the following results.

(a) $d(u, \mathcal{C}) = q - k - 2$ if (k, b, c) satisfies one of the following conditions.

- $2 \mid k+1, 4 \nmid k+1$, *and* $b^2 \neq c$.
- $2 \mid k+1, 4 \nmid k+1$, $b^2 = c$ *and* $k + 2 > q/2$.
- $4 \mid k+1$ *and* $c \neq 0$.
- $4 \mid k+1$,$c = 0$ *and* $k + 2 < q/2$.

(b) $d(u, \mathcal{C}) \leq q - k - 1$ if (k, b, c) satisfies one of the following conditions.

- $4 \mid k+1$, $c = 0$ *and* $k + 2 \geq q/2$.
- $4 \mid k$.
- $2 \mid k$,$4 \nmid k$, *and* $b \neq 0$.
- $2 \mid k$,$4 \nmid k$, *and* $c \neq 0$.
- $2 \mid k$,$4 \nmid k$, $b = c = 0$ *and* $k + 1 > q/2$.

(iii) If $p \neq 2$ and $k + 2 \leq q - 2$, then if $p \nmid k + 2$, we have $d(u, \mathcal{C}) \leq q - k - 1$. In the case that $p \mid k + 2$ we can conclude the following results.

(a) If $b = c = 0$ and $k + 2 > q/2 + 1$, then $d(u, \mathcal{C}) \leq q - k - 1$;

(b) If $b \neq 0$, then $d(u, \mathcal{C}) = q - k - 2$;

(c) If $c \neq 0$, then

$$d(u, \mathcal{C}) = \begin{cases} q - k & \textit{if } p = 3, k + 2 = 3 \textit{ and } -c \textit{ is not a nonzero square}, \\ q - k - 1 & \textit{if } p = 3, k + 2 = q - 3 \textit{ and } -c \textit{ is not a nonzero square}, \\ q - k - 2 & otherwise. \end{cases}$$

In particular, for the cases which do not satisfy the conditions we discuss in our theorem, we find some new deep holes.

- $q=8$, $k=1$, $b^2=c\in\mathbb{F}_8$. Received word with polynomial $u=x^3+bx^2+cx+d$ is a deep hole of $\mathcal{C}_8=\{(x_i,x_i,\dots,x_i)\in\mathbb{F}_8^8\mid x_i\in\mathbb{F}_8,1\le i\le 8\}$, where $d\in\mathbb{F}_8$.
- $q=8$, $k=2$, $b^2=c=0$.Received word with polynomial $u=x^4+dx+e$ is a deep hole of $\mathcal{C}_8=\{(mx_1+t,mx_2+t,\dots,mx_8+t)\in\mathbb{F}_8^8\mid x_i\in\mathbb{F}_8,1\le i\le 8,m,t\in\mathbb{F}_8\}$, where $d,e\in\mathbb{F}_8$.

In our proof, we convert the problem of deciding the error distance of a received word to solving a polynomial equation. Compared with approach in [2][3], our method is much simpler and using some algebraic constructions and character sum estimate, we not only get the deep hole results, but also can determine the error distance explicitly.

Organization. In Section 2, we provide a brief introduction to finite fields and state some fundamental definitions and lemmas. In Section 3, we discuss the error distance of received words of degree $k+1$ to Standard Reed-Solomon code and Primitive Reed-Solomon code respectively. The case that computing error distance of received words of degree $k+2$ to Standard Reed-Solomon code is studied in Section 4.

2 Preliminaries

We first review the theory of finite field and character sums in the form we need. Let $\mathbb{F}_q$ be the finite field with character p, where q is a p power. For a element $a\in\mathbb{F}_q^*$, the order of a is defined by the smallest number d such that $a^d=1$. Let $\chi:\mathbb{F}_q^*\longrightarrow\mathbb{C}^*$ be a multiplicative character from the invertible elements of $\mathbb{F}_q$ to the non-zero complex numbers and satisfies that for $a,b\in\mathbb{F}_q^*$, $\chi(ab)=\chi(a)\chi(b)$. If for all $a\in\mathbb{F}_q^*$, $\chi(a)=1$, then call χ trivial character, denoted by 1. The smallest d such that $\chi^d=1$ is called the degree of χ. Extend the definition to $\mathbb{F}_q$ by

$$\chi(0)=\begin{cases}1, & \chi=1;\\ 0, & \chi\neq 1\end{cases}$$

Lemma 2 *[9] Let $\mathbb{F}_q$ be a finite field, $p\neq 2$. If n is odd and $a_i\neq 0, 1\le i\le n$, then the number of solutions of equation $a_1x_1^2+\cdots+a_nx_n^2=b$ over $\mathbb{F}_q$ is*

$$q^{n-1}+q^{(n-1)/2}\eta((-1)^{(n-1)/2}ba_1\cdots a_n),$$

where η is a character of degree 2 over $\mathbb{F}_q$.

Lemma 3 *[9]Let $\mathbb{F}_q$ be a finite field, $p\neq 2$. If n is even and $a_i\neq 0, 1\le i\le n$, then the number of solutions of equation $a_1x_1^2+\cdots+a_nx_n^2=b$ over $\mathbb{F}_q$ is*

$$q^{n-1}+v(b)q^{(n-2)/2}\eta((-1)^{n/2}a_1\cdots a_n),$$

where η is a character of degree 2 over $\mathbb{F}_q$.

Lemma 4 *[10] Let $u \in \mathbb{F}_q^n$ be a received word with degree $k+r$, where $k+1 \leq k+r \leq n-1$. Then*

(i) $d(u, \mathcal{C}) = n-k-r$ *if and only if there exists a subset $E = \{x_1, \ldots, x_{k+r}\}$ of $\mathcal{D}$ such that*

$$u(x) - v(x) = (x - x_1) \cdots (x - x_{k+r}),$$

for some $v(x) \in \mathbb{F}_q[x],\ \deg v(x) \leq k-1$.

(ii) $d(u, \mathcal{C}) \leq n-k-i,\ (1 \leq i \leq r)$ *if and only if there exists a subset $E = \{x_1, \ldots, x_{k+i}\}$ of $\mathcal{D}$ and a monic polynomial$g(x)$ of degree $r-i$ such that*

$$u(x) - v(x) = (x - x_1) \cdots (x - x_{k+i}) g(x),$$

for some $v(x) \in \mathbb{F}_q[x],\ \deg v(x) \leq k-1$.

3 The case for received words of degree $k+1$

In this section, we give the proof of Theorem 1. Let $u \in \mathbb{F}_q^n$ be a received word represented by polynomial $u(x) = x^{k+1} - bx^k + v(x)$ with $\deg(v) \leq k-1$.

3.1 Computing $d(u, \mathcal{C}_q)$

From Lemma 1 and Lemma 4, we know that $q-k-1 \leq d(u, \mathcal{C}_q) \leq q-k$, and $d(u, \mathcal{C}_q) = q-k-1$ if and only if there exists a subset $\{x_1, x_2, \ldots, x_{k+1}\} \subset \mathbb{F}_q$ of size $k+1$ such that $b = x_1 + \cdots + x_{k+1}$.

- $b \neq 0$.

  Let g be a primitive element in $\mathbb{F}_q$, and

  $$u = 1 + g + g^2 + \cdots + g^{k-1} + 0.$$

  as the order of g is $q-1$ and $k+1 \leq q-1$, then $u \neq 0$ and

  $$1 = 0 + u^{-1} + u^{-1}g + \cdots + u^{-1}g^{k-1},$$

  thus,

  $$b = 0 + bu^{-1} + bu^{-1}g + \cdots + bu^{-1}g^{k-1},$$

  obviously, the $k+1$ items above are distinct. To be concluded, if $b \neq 0$, $d(u, \mathcal{C}_q) = q-k-1$.

- $b = 0$.

- $b = 0, p \neq 2$.

  In this case, for any $x \in \mathbb{F}_q^*$, $x \neq -x$, therefore

  $$\mathbb{F}_q = \{0, x_1, -x_1, \ldots, x_{\frac{q-1}{2}}, -x_{\frac{q-1}{2}}\}.$$

  If k is odd, then

  $$0 = x_1 + (-x_1) + \cdots + x_{\frac{k+1}{2}} + (-x_{\frac{k+1}{2}}).$$

  If k is even, we only need to add 0 to the right side of the equation above. Therefore, if $b = 0, p \neq 2$, $d(u, \mathcal{C}_q) = q - k - 1$.

- $b = 0, p = 2$.

  Without loss of generality, we can assume $q > 2$. As the sum of all the elements in $\mathbb{F}_q$ is 0, the conclusion holds for $k + 1$ iff it holds for $q - k - 1$. For $k + 1 = 2$, $d(u, \mathcal{C}_q) = q - k - 1$ is equivalent to the fact that there exists $x_1, x_2 \in \mathbb{F}_q$, $x_1 \neq x_2$ and $x_1 + x_2 = 0$. But when $p = 2$, if $x_1 + x_2 = 0$, then $x_1 = x_2$, contradiction. Thus, if $p = 2, b = 0$, and $k + 1 = 2$, $d(u, \mathcal{C}_q) = q - k$. Likewise, if $p = 2, b = 0$, and $q - k - 1 = 2$, $d(u, \mathcal{C}_q) = q - k$. Without loss of generality, we can assume $2 < k + 1 \leq q/2$. Set

  $$S = \mathbb{F}_q^* \setminus \{g, g^2, \ldots, g^{k-1}\}.$$

  It is easy to see that the number of elements in S is $q - 1 - (k - 1) = q - k > q/2$. Set

  $$T = \{g + g^2 + \cdots + g^{k-1} + g^i \mid g^i \in S\},$$

  then the number of elements in T is also $q - k$, thus, $|S| + |T| > q$, and $S \cup T \subseteq \mathbb{F}_q$, which means that there exist two elements g^i and g^j in S such that

  $$g + g^2 + \cdots + g^{k-1} + g^i = g^j.$$

  As $p = 2$, then we have

  $$g + g^2 + \cdots + g^{k-1} + g^i + g^j = 0.$$

  Obviously, these $k + 1$ elements are distinct, so far, we can conclude that $d(u, \mathcal{C}_q) = q - k - 1$. Likewise, the same conclusion holds for $2 < q - k - 1 \leq q/2$. Overall, if $p = 2, b = 0$ and $1 < k < q - 3$, then $d(u, \mathcal{C}_q) = q - k - 1$. As for the case $k = q - 2$, for any b, the sum of $q - 1$ elements in $\mathbb{F}_q \setminus \{-b\}$ is b. Thus, if $k = q - 2$, $d(u, \mathcal{C}_q) = q - k - 1$. The proof of the first part of Theorem 1 is complete.

3.2 Computing $d(u, \mathcal{C}_q^*)$

The proof of Theorem 1(ii) is similar to the proof of Theorem 1(i).

- $b \neq 0$. For this case, the proof is the same as the proof for $b \neq 0$ in section 3.1. We omit it and conclude that if $b \neq 0$, $d(u, \mathcal{C}_q^*) = q - k - 2$.
- $b = 0$.
  - $b = 0, k + 1 = q - 2$.

    If there exist $k + 1$ distinct elements $x_1, x_2, \ldots, x_{k+1}$ in $\mathbb{F}_q^*$ satisfying $\sum x_i = 0$, there is only one nonzero element in $\mathbb{F}_q^* \setminus \{x_1, x_2, \ldots, x_{k+1}\}$, which contradicts the fact that the sum of all elements in $\mathbb{F}_q^*$ is 0. Thus, if $b = 0$ and $k + 1 = q - 2$, $d(u, \mathcal{C}) = q - k - 1$.
  - $b = 0, k + 1 \neq q - 2$.

    If $p \neq 2$, and $k + 1$ is even, the proof is same as the proof for the case $b = 0, p \neq 2$, and $k + 1$ is even in Section 3.1. Now we discuss the case that $k+1$ is odd. As $q > 5$, we can find $z_1, z_2 \in \mathbb{F}_q^*$ satisfying $z_1 \neq z_2, -z_2, -2z_2, -\frac{1}{2}z_2$. Thus,

    $$\mathbb{F}_q = \{z_1 + z_2 + z | z \in \mathbb{F}_q\}.$$

    So there exists $z_3 \in \mathbb{F}_q^*$ such that z_1, z_2, z_3 are distinct and

    $$z_1 + z_2 + z_3 = 0.$$

    As the sum of all the elements in $\mathbb{F}_q^*$ is 0, the conclusion holds for $k+1$ iff it holds for $q-k-2$. As $k + 1$ is odd and $k + 1 \neq q - 2$, we can assume $k + 1 \leq q - 4$, say $k - 2 \leq q - 7$. Set

    $$M = \mathbb{F}_q^* \setminus \{\pm z_1, \pm z_2, \pm z_3\}$$

    then the number of elements in M is also $q - 7$, together with the fact that $k - 2$ is even, we can get $k - 2$ elements in M summing to 0 similarly to what we proved in last subsection. Adding z_1, z_2, z_3 into these $k - 2$ elements, then we get $k + 1$ distinct elements in $\mathbb{F}_q^*$ whose sum is 0. If $p = 2$, the proof is same as the proof for the case $p = 2, b = 0$ in section 3.1.

This completes the proof of Theorem 1.

4 The case for received words of degree $k + 2$

Lemma 5 *(i) Suppose $p = 2$, $2 \leq t \leq q-2$, and $c \in \mathbb{F}_q^*$. Then there exist t distinct elements $\gamma_1, \gamma_2, \ldots, \gamma_t$ in $\mathbb{F}_q^*$ such that*

$$c = \sum_{1 \leq i < j \leq t} \gamma_i \gamma_j.$$

(ii) *Suppose* $p = 2$, $2 \le t \le q-3$, *and* $c \in \mathbb{F}_q^*$, *then there exist* t *distinct elements* $\gamma_1, \gamma_2, \dots, \gamma_t$ *in* $\mathbb{F}_q^*$ *such that*

$$c = \sum_{1 \le i \le j \le t} \gamma_i \gamma_j .$$

(iii) *If* $t = q-1$, *then* $\sum_{\substack{1 \le i < j \le q-1 \\ \gamma_i, \gamma_j \in \mathbb{F}_q^*}} \gamma_i \gamma_j = 0.$

Proof. Let g be a primitive element in $\mathbb{F}_q$.

(i) As $2 \le t \le q-2$, $(1-g^{t-1})(1-g^t) \ne 0$. Therefore, for $p = 2$ and any $c \in \mathbb{F}_q^*$, the following equation with variable y always has solutions.

$$y^2 \frac{1}{1-g} \frac{g(1-g^{t-1})(1-g^t)}{1-g^2} = c.$$

Suppose σ is a root of the equation above and set $\gamma_i = \sigma g^{i-1}, 1 \le i \le t$, then

$$\begin{aligned}
\sum_{1 \le i < j \le t} \gamma_i \gamma_j &= \sigma^2 \sum_{0 \le i < j \le t-1} g^{i+j} \\
&= \sigma^2 [(g + g^2 + \cdots + g^{t-1}) + g(g^2 + \cdots + g^{t-1}) + \cdots \\
&+ g^{t-4}(g^{t-3} + g^{t-2} + g^{t-1}) + g^{t-3}(g^{t-2} + g^{t-1}) + g^{t-2} g^{t-1}] \\
&= \sigma^2 \frac{1}{1-g} \{(g + g^3 + \cdots + g^{2t-7} + g^{2t-5} + g^{2t-3}) \\
&- (g^t + g^{t+1} + \cdots + g^{2t-4} + g^{2t-3} + g^{2t-2})\} \\
&= \sigma^2 \frac{1}{1-g} \frac{g(1-g^{t-1})(1-g^t)}{1-g^2} \\
&= c.
\end{aligned}$$

(ii) According to $2 \le t \le q-3$, we can conclude that $(1-g^{t+1})(1-g^t) \ne 0$. Then for any $c \in \mathbb{F}_q^*$, let ω be a solution of the following equation.

$$y^2 \frac{1}{1-g} \frac{(1-g^{t+1})(1-g^t)}{1-g^2} = c.$$

Set $\gamma_i = \omega g^{i-1}, 1 \le i \le t$. Similarly to (i), we can deduce that

$$\begin{aligned}
\sum_{1 \le i < j \le t} \gamma_i \gamma_j &= \omega^2 \sum_{0 \le i \le j \le t-1} g^i g^j \\
&= \omega^2 \{1 + g^2 + \cdots + g^{2(t-1)} + \sum_{0 \le i < j \le t-1} g^{i+j}\} \\
&= \omega^2 \frac{1}{1-g} \frac{(1-g^{t+1})(1-g^t)}{1-g^2} \\
&= c.
\end{aligned}$$

(iii) For the case that $t = q-1$, all the elements in $\mathbb{F}_q^*$ have the form $g^i,\ i = 1, 2, \ldots, q-2$, so

$$\begin{aligned}\sum_{0\le i<j\le q-2} g^i g^j &= \sum_{0\le i<j\le q-2} g^{i+j} \\ &= \frac{1}{1-g}\frac{g(1-g^{q-2})(1-g^{q-1})}{1-g^2} \\ &= 0\end{aligned}$$

Lemma 6 *Assume that $1 \le t < \frac{q}{2} - 1$. If $p = 2$, suppose $4 \mid t$ and if $p \ne 2$, suppose $p \mid t$, then there exist t distinct elements $\xi_1, \ldots, \xi_t$ in $\mathbb{F}_q^*$ such that*

$$\sum_{1\le i<j\le t} \xi_i \xi_j = 0.$$

Proof. Let g be a primitive element in $\mathbb{F}_q$, set

$$M_1 = \frac{1-g^{t-1}}{1-g}, M_2 = \frac{1}{1-g}\frac{g(1-g^{t-2})(1-g^{t-1})}{1-g^2}.$$

Let $\phi \in \mathbb{F}_q^*$ satisfy $\phi \ne -M_1, \phi^2 + 2\phi M_1 + M_2 \ne 0$, for $2 \le i \le t$, $\phi \ne g^{i-2}$ and if $g^{i-2} + M_1 \ne 0$,

$$\phi \ne -\frac{M_2 + g^{i-2}M_1}{g^{i-2} + M_1}.$$

As $1 \le t < \frac{q}{2} - 1$,

$$(q-1) - 1 - 2 - 2(t-1) = q - 2 - 2t > 0.$$

So ϕ does exist. Set

$$z = \frac{\phi M_1 + M_2}{\phi + M_1}$$

and $\xi_1 = z + \phi,\ \xi_i = z + g^{i-2}, 2 \le i \le t$. Because of the choice of ϕ, we can easily varify that $\xi_1 \ne 0$, and $\xi_i \ne \xi_j$, $i \ne j$. For $2 \le i \le t$, if $g^{i-2} + M_1 \ne 0$, obviously, $z + g^{i-2} \ne 0$, i.e. $\xi_i \ne 0$. If $g^{i-2} + M_1 = 0$,

$$\begin{aligned}\xi_i &= z + g^{i-2} \\ &= \frac{\phi M_1 + M_2}{\phi + M_1} + g^{i-2} \\ &= \frac{\phi M_1 + M_2}{\phi + M_1} - M_1 \\ &= \frac{M_2 - M_1^2}{\phi + M_1}.\end{aligned}$$

We have

$$\begin{aligned}M_2 - M_1^2 &= \frac{1}{1-g}\frac{g(1-g^{t-2})(1-g^{t-1})}{1-g^2} - \frac{(1-g^{t-1})^2}{(1-g)^2} \\ &= \frac{(1-g^{t-1})(g^t - 1)}{(1-g)^2(1+g)} \\ &\ne 0.\end{aligned}$$

Then $\xi_i \neq 0$. Under the condition that $p = 2$, $4 \mid t$, or $p \neq 2$, $p \mid t$, we can get

$$\begin{aligned}
\sum_{1\le i<j\le t} \xi_i\xi_j &= (z+\phi)\sum_{2\le i\le t}(z+g^{i-2}) + \sum_{2\le i<j\le t}(z+g^{i-2})(z+g^{j-2}) \\
&= \frac{t(t-1)}{2}z^2 + (t-1)(\phi+1+g+\cdots+g^{t-2})z + \phi M_1 + \sum_{2\le i<j\le t} g^{i+j-4} \\
&= -(\phi+M_1)z + \phi M_1 + M_2 \\
&= 0.
\end{aligned}$$

By lemma 5 (iii) and lemma 6, we can easily get the following corollary.

Corollary 1 *Assume that $t > \frac{q}{2}$. If $p = 2$, suppose $4 \mid q-1-t$ and if $p \neq 2$, suppose $p \mid q-1-t$, then there exist t distinct elements $\xi_1, \ldots, \xi_t$ in $\mathbb{F}_q^*$ such that*

$$\sum_{1\le i\le j\le t} \xi_i\xi_j = 0.$$

Lemma 7 *Assume that $p \neq 2$, $r, r_1, \mu \neq 0$, and $b, c \in \mathbb{F}_q$, denote $A = \left\{\frac{\alpha^2}{\mu} - \frac{b^2r^2}{\mu} + 2cr_1 \mid \alpha \in \mathbb{F}_q\right\}$. If $2 < t < \frac{q+1}{2}$ and t is even, then there exist t distinct elements $y_1, \cdots, y_t$ in $\mathbb{F}_q^*$ such that*

$$(y_1+\cdots+y_t)^2 - r(y_1^2+\cdots+y_t^2) \in A.$$

Proof. Let g be a primitive element of $\mathbb{F}_q$. From Lemma 2, the following equation with variables α, y, z has at most (q^2-1) nonzero solutions in $\mathbb{F}_q^3$.

$$\alpha^2 + g^2z^2\frac{1-g^{t-2}}{1-g^2} - g^2y^2\frac{1-g^t}{1-g^2} = 0 \tag{1}$$

Denote

$T = \{(\alpha, z) \in \mathbb{F}_q^2 \mid \alpha, z \neq 0,$ and there exists $y \in \mathbb{F}_q^*$ such that $(\alpha, z, y)$is a nonzero solution of Equation (1)$\}$.

Thus, there are at most $\frac{q^2-1}{2}$ elements in T. The number of pairs $(\alpha, z) \in \mathbb{F}_q^2$ such that $\alpha \neq 0$ and $\alpha \neq \pm zg^i, 1 \le i \le (t-2)/2, z \neq 0$ is $(q-1)(q-t+1)$. For $\frac{q+1}{2} > t > 2$,

$$(q-1)(q-t+1) - \frac{q^2-1}{2} > 0,$$

then there exist $\alpha_1, z_1 \in \mathbb{F}_q^*, \alpha_1 \neq \pm z_1g^i, 1 \le i \le (t-2)/2$, and for all $y \in \mathbb{F}_q^*$

$${\alpha_1}^2(-2r) + (-2r)g^2z^2\frac{1-g^{t-2}}{1-g^2} \neq (-2r)g^2y^2\frac{1-g^t}{1-g^2}. \tag{2}$$

Set $y_1 = yg, y_2 = -yg, \ldots, y_{t-1} = yg^{\frac{t}{2}}, y_t = -yg^{\frac{t}{2}}, y \in \mathbb{F}_q^*, m = y_1+y_2+\cdots+y_t, n = y_1^2+y_2^2+\cdots+y_t^2$, we have

$$m^2 - rn = (-2r)g^2y^2\frac{1-g^t}{1-g^2}.$$

Set $y_1 = \alpha_1, y_2 = -\alpha_1, y_3 = z_1 g, y_4 = -z_1 g, \ldots, y_{t-1} = z_1 g^{\frac{t-2}{2}}, y_t = -z_1 g^{\frac{t-2}{2}}, m = y_1 + y_2 + \cdots + y_t$, $n = y_1^2 + y_2^2 + \cdots + y_t^2$, then

$$m^2 - rn = (-2r){\alpha_1}^2 + (-2r)g^2 {z_1}^2 \frac{1-g^{t-2}}{1-g^2}.$$

According to (2), denote

$$B = \left\{(-2r)g^2y^2\frac{1-g^t}{1-g^2} \mid y \in \mathbb{F}_q^*\right\} \cup \left\{(-2r)({\alpha_1}^2 + g^2{z_1}^2\frac{1-g^{t-2}}{1-g^2})\right\}.$$

Therefore, $|A| + |B| = q + 1$, which implies that $A \cap B \neq \emptyset$. In other words, there exist t distinct elements $y_1, \cdots, y_t$ in $\mathbb{F}_q^*$ such that $m^2 - rn \in A$.

4.1 Proof of theorem 2 (i).

i) $b^2 = c$.

From Lemma 4, we know that $d(u, \mathcal{C}) = q - k - 2$ iff there are $k + 2$ distinct elements $x_1, \ldots, x_{k+2}$ in $\mathbb{F}_q$ satisfying

$$\begin{cases} b = x_1 + \cdots + x_{k+2}, \\ c = \sum_{1 \le i < j \le k+2} x_i x_j. \end{cases} \tag{3}$$

Denote $\mathbb{F}_q^* = \{a_1, \ldots, a_{q-1}\}$.

- $b = 0, c = 0$.

  Set $x_i = a_i, 1 \le i \le q-1$. The conclusion holds because of the fact that all elements in $\mathbb{F}_q^*$ sum to 0 and Lemma 5.

- $b^2 = c, b \neq 0$.

  Without loss of generality, assume that $b = -a_1$, and set $x_1 = 0, x_2 = a_2, \cdots, x_{q-1} = a_{q-1}$, then $b = -a_1 = a_2 + \cdots + a_{q-1} = 0 + x_2 + \cdots + x_{q-1}$. From Lemma 5,

$$\begin{aligned} c &= 0 + b^2 \\ &= 0 + b(a_2 + \cdots + a_{q-1}) \\ &= \{\sum_{1 \le i < j \le q-1} a_i a_j\} - a_1(a_2 + \cdots + a_{q-1}) \\ &= \sum_{2 \le i < j \le q-1} a_i a_j \\ &= \sum_{1 \le i < j \le q-1} x_i x_j. \end{aligned}$$

  Here we complete the proof that when $b^2 = c$ and $k + 2 = q - 1$, $d(u, \mathcal{C}) = q - k - 2$.

ii) $b^2 \neq c$.

For this case, if we prove that $d(u, \mathcal{C}) \neq q-k-2$ and $d(u, \mathcal{C}) \leq q-k-1$, then by lemma 1 we have $d(u, \mathcal{C}) = q-k-1$. Firstly, we prove that $d(u, \mathcal{C}) \neq q-k-2$.

- $b^2 \neq c, b = 0$.

  If there are $q-1$ distinct elements $x_1, \cdots, x_{q-1}$ in $\mathbb{F}_q$ such that $b = 0 = x_1 + \cdots + x_{q-1}$. As all the nonzero elements in $\mathbb{F}_q$ sum to 0, then $x_i \neq 0, 1 \leq i \leq q-1$. From Lemma 5, $c = 0$, which is contradiction with the fact that $b^2 \neq c$.

- $b^2 \neq c, b \neq 0$.

  If there are $q-1$ distinct elements $x_1, \cdots, x_{q-1}$ in $\mathbb{F}_q$ such that $b = x_1 + \cdots + x_{q-1}$. As all the nonzero elements in $\mathbb{F}_q$ sum to 0, then there exists $x_i = 0$. Assume that $x_1 = 0$, $x_2 = a_2, \cdots, x_{q-1} = a_{q-1}$, then $b = a_2 + \cdots + a_{q-1} = -a_1$.

$$\begin{aligned} c &= \sum_{1 \leq i < j \leq q-1} x_i x_j \\ &= \sum_{2 \leq i < j \leq q-1} a_i a_j \\ &= \sum_{1 \leq i < j \leq q-1} a_i a_j - a_1(a_2 + \cdots + a_{q-1}) \\ &= a_1^2 \\ &= b^2. \end{aligned}$$

  Which is contradiction with $b^2 \neq c$.

Secondly, we prove that if $b^2 \neq c$, then $d(u, \mathcal{C}) \leq q-k-1$. From Lemma 4, we know that $d(u, \mathcal{C}) \leq q-k-1$ iff there are $k+1$ distinct elements $x_1, \ldots, x_{k+1}$ in $\mathbb{F}_q$ and $a \in \mathbb{F}_q$ satisfying

$$\begin{cases} b = x_1 + \cdots + x_{k+1} + a, \\ c = a(x_1 + \cdots + x_{k+1}) + \sum_{1 \leq i < j \leq k+1} x_i x_j. \end{cases} \tag{4}$$

As $b^2 \neq c$, then $b \neq 0$ or $c \neq 0$. If $b \neq 0$ and $c \neq 0$, set $\eta_1 = -b^{-1}c, \eta_2 = 0$. If $b \neq 0$ and $c = 0$, set $\eta_1 \neq 0, -b, -2b$ and $\eta_2 = -(b+\eta_1)^{-1} b \eta_1$. If $b = 0$ and $c \neq 0$, set $\eta_1 \neq 0, \eta_1^2 \neq -c$ and $\eta_2 = -\eta_1^{-1} c$. Then for each case we can check that $\eta_1 \neq \eta_2$. Let $x_i \in \mathbb{F}_q \setminus \{\eta_1, \eta_2\}, 1 \leq i \leq q-2, a = b + \eta_1 + \eta_2$. thus, for each case, we all have that

$$x_1 + \cdots + x_{q-2} + a = x_1 + x_2 + \cdots + x_{q-2} + b + \eta_1 + \eta_2 = b.$$

$$\begin{aligned} a(x_1+\cdots+x_{q-2})+\sum_{1\le i<j\le q-2} x_ix_j &= (b+\eta_1+\eta_2)(-\eta_1-\eta_2)-\eta_1(x_1+\cdots+x_{q-2}) \\ &- \eta_2(x_1+\cdots+x_{q-2}+\eta_1) \\ &= (-b-\eta_1)\eta_2-b\eta_1 \\ &= c. \end{aligned}$$

4.2 Proof of theorem 2 (ii).

(1) From Lemma 4, we know that $d(u,\mathcal{C})=q-k-2$ iff there are $k+2$ distinct elements $x_1,\ldots,x_{k+2}$ in $\mathbb{F}_q$ satisfying

$$\begin{cases} b=x_1+\cdots+x_{k+2}, \\ c=\sum_{1\le i<j\le k+2} x_ix_j. \end{cases} \tag{5}$$

Set $x_1=x+y_1,\cdots,x_{k+2}=x+y_{k+2}, m_1=y_1+\cdots+y_{k+2}, m_2=\sum_{1\le i<j\le k+2} y_iy_j$. Then

$$\begin{cases} b=(k+2)x+m_1 \\ c=\frac{(k+2)(k+1)}{2}x^2+(k+1)m_1x+m_2. \end{cases} \tag{6}$$

Obviously, $d(u,\mathcal{C})=q-k-2$ iff the equation above has a solution.

- $2|k+1$, $4\nmid k+1$.

  In this case, Equation (6) can be simplified as the following equation with variable x.

$$\begin{cases} b=x+m_1 \\ c=x^2+m_2. \end{cases} \tag{7}$$

  Equation (6) having a solution is equivalent to the fact that there are $k+2$ distinct elements $y_1,\cdots,y_{k+2}$ in $\mathbb{F}_q$ such that $b^2+c=\sum_{1\le i\le j\le k+2} y_iy_j$. If $b^2\ne c$, as $k+2\le q-2$ and $2|k+1$, then $k+2\le q-3$. From Lemma 5 there always exist $k+2$ distinct elements $\gamma_1,\gamma_2,\cdots,\gamma_{k+2}$ in $\mathbb{F}_q^*$ such that

$$b^2+c=\sum_{1\le i\le j\le k+2}\gamma_i\gamma_j.$$

  If $b^2=c$ and $k+2>q/2$, denote $t=q-1-(k+2)$. As $p=2, 2|k+1$, $4\nmid k+1$, so $4\mid t$ and $t<q/2-1$. From Lemma 6, we can get t distinct elements $\xi_1,\ldots,\xi_t$ in $\mathbb{F}_q^*$ such that

$$\sum_{1\le i<j\le t}\xi_i\xi_j=0.$$

  Denote $\mathbb{F}_q^*=\{\xi_1,\ldots,\xi_t,\xi_{t+1},\ldots,\xi_{q-1}\}$, using Lemma 5 and the fact that all the elements in $\mathbb{F}_q^*$

sum to 0, we can obtain

$$\begin{aligned} 0 &= \sum_{1\le i<j\le q-1} \xi_i\xi_j \\ &= \sum_{1\le i<j\le t} \xi_i\xi_j + (\xi_1+\cdots+\xi_t)(\xi_{t+1}+\cdots+\xi_{q-1}) + \sum_{t+1\le i<j\le q-1} \xi_i\xi_j \\ &= 0 + \sum_{t+1\le i\le j\le q-1} \xi_i\xi_j. \end{aligned}$$

Therefore, in the condition that $2 \mid k+1$, $4 \nmid k+1$, if $b^2 \neq c$ or $b^2 = c$, $k+2 > q/2$, then $d(u,\mathcal{C}) = q-k-2$.

- $4|k+1$. We can simplify Equation (6) as

$$\begin{cases} b = x + m_1; \\ c = m_2. \end{cases} \tag{8}$$

Equation (8) having a solution is equivalent to the fact that there are $k+2$ distinct elements $y_1, \cdots, y_{k+2}$ in $\mathbb{F}_q$ such that $c = \sum_{1\le i<j\le k+2} y_i y_j$. If $c \neq 0$, the claim holds directly from Lemma 5. If $c = 0$, and $k+2 < q/2$, From Lemma 6, we can get $k+1$ distinct elements $\xi_1, \ldots, \xi_{k+1}$ in $\mathbb{F}_q^*$ such that

$$\sum_{1\le i<j\le k+1} \xi_i\xi_j = 0.$$

Thus, we can get $k+2$ distinct elements $0, \xi_1, \ldots, \xi_{k+1}$ in $\mathbb{F}_q$ sum to 0. Overall, when $4|k+1$, if $c \neq 0$ or $c = 0$, $k+2 < q/2$, then $d(u,\mathcal{C}) = q-k-2$.

(2) From Lemma 4, we know that $d(u,\mathcal{C}) \le q-k-1$ iff there are $k+1$ distinct elements $x_1, \ldots, x_{k+1}$ in $\mathbb{F}_q$ and $a \in \mathbb{F}_q$ satisfying

$$\begin{cases} b = x_1 + \cdots + x_{k+1} + a \\ c = a(x_1 + \cdots + x_{k+1}) + \sum_{1\le i<j\le k+1} x_i x_j. \end{cases}$$

Denote $x_1 = x + y_1, \cdots, x_{k+1} = x + y_{k+1}, m_1 = y_1 + \cdots + y_{k+1}, m_2 = \sum_{1\le i<j\le k+1} y_i y_j$. Then,

$$\begin{cases} b = (k+1)x + a + m_1 \\ c = \frac{(k+1)k}{2}x^2 + km_1x + (k+1)ax + am_1 + m_2. \end{cases}$$

Solving the equation system above is equivalent to solving the following equation with variable x.

$$-\frac{(k+1)(k+2)}{2}x^2 + x((k+1)b - (k+2)m_1) + bm_1 - m_1^2 + m_2 - c = 0. \tag{9}$$

Therefore, the problem boils down to deciding if there exist $k+1$ distinct elements $y_1, \ldots, y_{k+1}$ in $\mathbb{F}_q$ such that the equation (9) has a solution.

- $4|k$. In this case, Equation (9) can be reduced to

$$(x+m_1)^2+b(x+m_1)+m_2+c=0. \tag{10}$$

If $c\neq 0$, we can get $k+1$ distinct elements $\{\beta_1,\ldots,\beta_{k+1}\}\subset\mathbb{F}_q^*$ such that $c=\sum_{1\le i<j\le k+1}\beta_i\beta_j$ according to Lemma 5. Then, set $y_i=\beta_i$, $x=\beta_1+\cdots+\beta_{k+1}$ is a solution of Equation (10).

If $c=0$, Denote $\alpha\in\mathbb{F}_q^*$ and $\alpha\neq b$. According to Lemma 5, we can get $k+1$ distinct elements $\nu_1,\ldots,\nu_{k+1}$ in $\mathbb{F}_q^*$ such that $\alpha^2+b\alpha=\sum_{1\le i<j\le k+1}\nu_i\nu_j$. Then, set $y_i=\nu_i, 1\le i\le k+1$, $x=\alpha+\nu_1+\cdots+\nu_{k+1}$ is a solution of Equation (10).

- $2|k$, $4\nmid k$. In this case, Equation (9) can be reduced to

$$b(x+m_1)+m_1^2+m_2+c=0. \tag{11}$$

If $b\neq 0$, it is easy to see that Equation (11)has a solution.

If $b=0, c\neq 0$, Equation (11)holding is equivalent to $m_1^2+m_2+c=0$, which can be deduced directly for Lemma 5.

If $b=0, c=0$, Equation (11)holding is equivalent to $m_1^2+m_2=0$. Denote $t=q-1-(k+1)$. If $k+1>q/2$, then $4\mid t$,and $t<q/2-1$. From Lemma 6, there exist t distinct $\xi_1,\ldots,\xi_t$ in $\mathbb{F}_q^*$ satisfying

$$0=\sum_{1\le i<j\le t}\xi_i\xi_j.$$

Denote $\mathbb{F}_q^*=\{\xi_1,\ldots,\xi_t,\xi_{t+1},\ldots,\xi_{q-1}\}$. Similarly to our proof in the first part, we can get $\sum\limits_{t+1\le i\le j\le q-1}\xi_i\xi_j=0$.

- $4\mid k+1$.

In this case, equation (9) can be reduced to

$$m_1x+bm_1+m_1^2+m_2+c=0.$$

Obviously, this equation has a solution. So, if $4\mid k+1$, $d(u,\mathcal{C})\le q-k-1$.

4.3 Proof of theorem 2 (iii)

In order to prove the third part of Theorem 2, we have to prove the following lemmas.

Lemma 8 *If $p\neq 2$, $3<k+2<q-3$,$k+2\neq\frac{q-1}{2}$, $k+2\neq\frac{q+1}{2}$, then for any $\zeta\in\mathbb{F}_q^*$, there exist $k+2$ distinct elements $y_1,\cdots,y_{k+2}$ in $\mathbb{F}_q$ such that*

$$\begin{aligned} y_1+\cdots+y_{k+2} &= 0\\ y_1^2+\cdots+y_{k+2}^2 &= \zeta. \end{aligned}$$

Proof.

1) k is even.

If $2 < k+2 < \frac{q-1}{2}$, denote $y_1 = \alpha, y_2 = -\alpha, y_3 = y, y_4 = -y, y_5 = yg, y_6 = -yg, \cdots, y_{k+1} = yg^{\frac{k}{2}-1}, y_{k+2} = -yg^{\frac{k}{2}-1}$. From Lemma 3, the following equation with variables α, y has at least $q-1$ solutions in $\mathbb{F}_q^2$

$$2(\alpha^2 + y^2 \frac{1-g^k}{1-g^2}) - \zeta = 0 \tag{12}$$

If $\alpha = 0$ or $y = 0$ or $\alpha = \pm yg^i, 0 \le i \le \frac{k}{2} - 1$, the number of solutions in $\mathbb{F}_q^2$ is at most $2k+4$. When $k+2 < \frac{q-1}{2}$, $2k+4 < q-1$. Then there exists solution of equation (12) satisfying $\alpha \ne 0, y \ne 0, \alpha \ne \pm yg^i, 0 \le i \le \frac{k}{2} - 1$. Assume that it's (α_1, y_1). Thus, there exist $k+2$ nonzero elements $\alpha_1, -\alpha_1, y_1, -y_1, y_1 g, -y_1 g \cdots, y_1 g^{\frac{k}{2}-1}, -y_1 g^{\frac{k}{2}-1}$ in $\mathbb{F}_q$ such that

$$\begin{aligned}&\alpha_1 + (-\alpha_1) + y_1 + (-y_1) + y_1 g + (-y_1 g) + \cdots + y_1 g^{\frac{k}{2}-1} + (-y_1 g^{\frac{k}{2}-1}) = 0,\\ &\alpha_1^2 + (-\alpha_1)^2 + y_1^2 + (-y_1)^2 + \cdots + (y_1 g^{\frac{k}{2}-1})^2 + (-y_1 g^{\frac{k}{2}-1})^2 = 2(\alpha_1^2 + y_1^2 \tfrac{1-g^k}{1-g^2}) = \zeta.\end{aligned}$$

2) k is odd.

According to the first part of our proof, we know that if $2 < k+1 < \frac{q-1}{2}$, then there exist $k+1$ distinct elements $y_1, \cdots, y_{k+1}$ in $\mathbb{F}_q^*$ such that

$$\begin{aligned} y_1 + \cdots + y_{k+1} &= 0 \\ y_1^2 + \cdots + y_{k+1}^2 &= \zeta. \end{aligned}$$

Set $y_{k+2} = 0$, then the conclusion holds. So, if k is odd, and $3 < k+2 < \frac{q+1}{2}$, the conclusion holds.

3) If k is even and $\frac{q-1}{2} < k+2 < q-3$, denote $t = q-k-2$, then t is odd and $3 < t < \frac{q+1}{2}$. From 2), we can see that if $3 < t < \frac{q+1}{2}$, there exist t distinct elements $z_1, \cdots, z_t$ in $\mathbb{F}_q$ such that

$$\begin{aligned} z_1 + \cdots + z_t &= 0 \\ z_1^2 + \cdots + z_t^2 &= -\zeta. \end{aligned}$$

Note that all elements in $\mathbb{F}_q$ satisfy the following properties.

$$\begin{aligned} \sum_{x \in \mathbb{F}_q} x &= 0 \\ \sum_{x \in \mathbb{F}_q} x^2 &= 0 \end{aligned}$$

Then the conclusion holds.

If k is odd and $\frac{q+1}{2} < k+2 < q-3$, denote $t = q-k-2$, then t is even and $3 < t < \frac{q-1}{2}$. From 1), we can see that if $2 < t < \frac{q-1}{2}$, there exist t distinct elements $z_1, \cdots, z_t$ in $\mathbb{F}_q$ such that

$$\begin{aligned} z_1 + \cdots + z_t &= 0 \\ z_1^2 + \cdots + z_t^2 &= -\zeta. \end{aligned}$$

Similarly, we can prove the conclusion holds.

Hence, the proof is complete.

Corollary 2 *If $p \neq 2$, $3 < k+2 < q-3$, and $p \mid k+2$ then for any $\zeta \in \mathbb{F}_q^*$, there exist $k+2$ distinct elements $y_1, \cdots, y_{k+2}$ in $\mathbb{F}_q$ such that*

$$\begin{aligned} y_1 + \cdots + y_{k+2} &= 0 \\ y_1^2 + \cdots + y_{k+2}^2 &= \zeta. \end{aligned}$$

Proof of theorem 2 (iii).

From Lemma 4, $d(u, \mathcal{C}) = q-k-2$ iff there are $k+2$ distinct elements $x_1, \cdots, x_{k+2}$ in $\mathbb{F}_q$ such that

$$\begin{cases} b = x_1 + \cdots + x_{k+2}, \\ c = \sum_{1 \leq i < j \leq k+2} x_i x_j. \end{cases}$$

It is equivalent to

$$\begin{cases} b = x_1 + \cdots + x_{k+2}, \\ b^2 - 2c = x_1^2 + \cdots + x_{k+2}^2. \end{cases} \tag{13}$$

Denote $x_1 = x + y_1, \cdots, x_{k+2} = x + y_{k+2}, m = y_1 + \cdots + y_{k+2}, n = y_1^2 + \cdots + y_{k+2}^2$. In order to prove that equation (13) has a solution, we just need to prove that there exist $k+2$ distinct elements $y_1, \cdots, y_{k+2}$ in $\mathbb{F}_q$ such that the following equation holds.

$$\begin{cases} b = (k+2)x + m, \\ b^2 - 2c = (k+2)x^2 + 2mx + n. \end{cases} \tag{14}$$

If $p \mid k+2$ and $b \neq 0$, first according to Theorem 1, we can get $k+2$ distinct elements $\chi_1, \cdots, \chi_{k+2}$ in $\mathbb{F}_q$ such that $b = x_1 + \cdots + x_{k+2}$, then $x = \frac{b^2 - 2c - n}{2b}$ is a solution of equation (14).

If $p \mid k+2$ and $b = 0$, equation (14) having a solution is equivalent to the fact that there are $k+2$ distinct elements $y_1, \cdots, y_{k+2}$ in $\mathbb{F}_q$ such that the following equation system holds.

$$\begin{aligned} y_1 + \cdots + y_{k+2} &= 0 \\ y_1^2 + \cdots + y_{k+2}^2 + 2c &= 0 \end{aligned}$$

When $c \neq 0$ and $3 < k+2 < q-3$, it can be deduced directly from Corollary 2. If $k+2 = 3$ or $k+2 = q-3$, since $p \mid k+2$, then $p = 3$. For the case $p = 3$ and $k+2 = 3$, we know that $d(u, \mathcal{C}) \leq q-k-1 = q-2$ iff there are 2 distinct elements x_1, x_2 in $\mathbb{F}_q$ and $a \in \mathbb{F}_q$ satisfying

$$\begin{cases} b = x_1 + x_2 + a = 0 \\ c = a(x_1 + x_2) + x_1 x_2. \end{cases} \tag{15}$$

As $p = 3$, Equation(15) having a solution is equivalent to $-c = (x_1 - x_2)^2$ holding for distinct x_1, x_2. Then when $-c$ is not a nonzero square, $d(u, \mathcal{C}) = q - k$. Moreover, if $-c = (x_1 - x_2)^2$ and $x_1 \neq x_2$, then $a = -(x_1 + x_2) \neq x_1$(otherwise $x_2 = x_1$, contradiction), likewise, $a \neq x_2$. Therefore there are 3 distinct elements in $\mathbb{F}_q$ such that Equation (15) holds, then $d(u, \mathcal{C}) = q-k-2$. In this way, we find a new deep hole with degree $k+2$ for $p = 3$ and $k = 1$, which means that Cheng-Murray Conjecture doesn't hold for this special case. But $k = 1$ is not usually used in designing Reed-Solomon code in practise, so this conjecture still needs further study. For the case $p = 3$ and $k+2 = q-3$, we know that $d(u, \mathcal{C}) \leq q-k-1$ iff there are $k+1$ distinct elements $x_1, \ldots, x_{k+1}$ in $\mathbb{F}_q$ and $a \in \mathbb{F}_q$ satisfying

$$\begin{cases} 0 = x_1 + \cdots + x_{k+1} + a \\ c = a(x_1 + \cdots + x_{k+1}) + \sum_{1 \leq i < j \leq k+1} x_i x_j. \end{cases}$$

which is equivalent to $-c = \sum_{1 \leq i < j \leq 4} x_i x_j$ holding for distinct $x_i \in \mathbb{F}_q,\ 1 \leq i \leq 4$. Obviously, the equation has solutions when $q > 3$. Moreover, $d(u, \mathcal{C}) = q-k-2$ iff the following equation has a solution for distinct x_1, x_2, x_3

$$\begin{cases} 0 = x_1 + x_2 + x_3 \\ c = \sum_{1 \leq i \leq j \leq 3} x_i x_j \end{cases} \tag{16}$$

which is equivalent to $-c = (x_1 - x_2)^2$ as $p = 3$. Using a similar argument, we know that $d(u, \mathcal{C}) = q-k-2$ iff $-c$ is a nonzero square.

From Lemma 4, we know that $d(u, \mathcal{C}) \leq q-k-1$ iff there are $k+1$ distinct elements $x_1, \ldots, x_{k+1}$ in $\mathbb{F}_q$ and $a \in \mathbb{F}_q$ satisfying

$$\begin{cases} b = x_1 + \cdots + x_{k+1} + a \\ c = a(x_1 + \cdots + x_{k+1}) + \sum_{1 \leq i < j \leq k+1} x_i x_j. \end{cases}$$

Denote $k+1 \equiv r \mod p, x_1 = x + y_1, \cdots, x_{k+1} = x + y_{k+1}, m = y_1 + \cdots + y_{k+1}, n = y_1^2 + \cdots + y_{k+1}^2$. Then,

$$\begin{cases} b = (k+1)x + a + m \\ c = \frac{(k+1)k}{2} x^2 + kmx + (k+1)ax + am + \frac{m^2 - n}{2}. \end{cases}$$

Solving the equation system above is equivalent to solving the following equation with variable x.

$$(r + r^2)x^2 + x(2m + 2mr - 2br) + m^2 + n - 2bm + 2c = 0 \tag{17}$$

We discuss the solution of Equation (17) according to the following three cases.

(i) $r = 0$.

In this case, Equation (17) can be reduced to the form $2mx + m^2 + n - 2bm + 2c = 0$. Obviously, it has a solution.

(ii) $r = -1$.

In this case, $p \mid k+2$, thus, for $b \neq 0$ or $b = 0,\ c \neq 0,\ p \neq 3$ or $p = 3$, we have discussed due to the first part of our proof. For $b = 0$ and $c = 0$, denote $t = q - 1 - (k+1)$, then $p|t$. From Corollary 1, if $k + 1 > q/2$, there are $k+1$ distinct elements $\xi_1, \ldots, \xi_{k+1}$ in $\mathbb{F}_q^*$ such that

$$\sum_{1 \leq i \leq j \leq k+1} \xi_i \xi_j = 0.$$

Then Equation (17) has a solution.

(iii) $r \neq -1, 0$.

Denote $S = \{x^2 \mid x \in \mathbb{F}_q\}$. As we know, equation $ax^2 + bx + c = 0 (a \neq 0)$ over finite field $\mathbb{F}_q$ with characteristic $p \neq 2$ having a solution is equivalent to discriminant $D = b^2 - 4ac \in S$. For Equation (17), The determinant

$$\begin{aligned} D_1 &= (2m + 2mr - 2br)^2 - 4(r + r^2)(m^2 + n - 2bm + 2c) \\ &= 4[(1+r)(m^2 - rn - 2cr) + b^2 r^2]. \end{aligned}$$

If $2 < k + 1 < \frac{q+1}{2}$ and $k+1$ is even, denote

$$A_1 = \Big\{ \frac{\alpha^2}{1+r} - \frac{b^2 r^2}{1+r} + 2cr \mid \alpha \in \mathbb{F}_q \Big\}.$$

According to Lemma 7, we can obtain that there are $k+1$ distinct nonzero elements $y_1, \cdots, y_{k+1}$ in $\mathbb{F}_q$ satisfying

$$(m^2 - rn) \in A_1,$$

which is equivalent to $D_1 = 4[(1+r)(m^2 - rn - 2cr) + b^2 r^2] \in S$.

Similarly, if $k+1$ is odd, we can get k distinct nonzero elements $y_1, \cdots, y_k$ in $\mathbb{F}_q$ satisfying $(m^2 - rn) \in A_1$, then set $y_{k+1} = 0$.

If $q - 3 \geq k + 1 \geq \frac{q+1}{2}$, denote $t = q - k - 1, t \equiv s \mod p$, then $s \neq 0, 1$ and $1 < t \leq \frac{q-1}{2} < \frac{q+1}{2}$. Set $m_1 = y_1 + \cdots + y_t, n_1 = y_1^2 + \cdots + y_t^2$, when $s \neq 0, 1, 1 < t < \frac{q+1}{2}$, the determinant of the following equation with variable x

$$(s - s^2)x^2 + x(2m_1 - 2m_1 s - 2bs) - m_1^2 + n_1 - 2bm_1 - 2c = 0 \qquad (18)$$

is

$$\begin{aligned} D_2 &= (2m_1 - 2m_1 s - 2bs)^2 - 4(s - s^2)(-m_1^2 + n_1 - 2bm_1 - 2c) \\ &= 4[(1-s)(m_1^2 - sn_1 + 2cs) + b^2 s^2]. \end{aligned}$$

From Lemma 7 there exist t distinct elements in $\mathbb{F}_q$ such that

$$D_2 \in S.$$

Then Equation (16) has a solution in $\mathbb{F}_q$, which is equivalent to the fact than the following equation has a solution in $\mathbb{F}_q^{t+1}$.

$$\begin{cases} -b = x_1 + \cdots + x_t - a, \\ -b^2 + 2c = x_1^2 + \cdots + x_t^2 - a^2. \end{cases} \tag{19}$$

Our conclusion holds directly by the following property of finite fields. When $q > 3$,

$$\begin{cases} \sum\limits_{x \in \mathbb{F}_q} x = 0 \\ \sum\limits_{x \in \mathbb{F}_q} x^2 = 0 \end{cases}$$

Acknowledgement. We express our deep gratitude to professor Daqing Wan for his useful suggestions. And we also acknowledge with gratitude the support by National Development Foundation for Cryptological Research(No. MMJJ201401003).

References

[1] Berlekamp, E., Welch, L.: Error correction of algebraic block codes. U.S. Patent Number 4633470, 1986.

[2] Cafure, A., Matera, G., Privitelli, M.: Singularities of symmetric hypersurfaces and an application to Reed-Solomon codes. Advances in Mathematics of Communications, 6(1):69–94, 2012.

[3] Cheng, Q., Li, J., Zhuang, J.: On Determining Deep Holes of Generalized Reed-Solomon Codes. `http://arxiv.org/pdf/1309.3546.pdf`, 2013.

[4] Cheng, Q., Murray, E. On deciding deep holes of Reed-Solomon codes. In Proceedings of TAMC 2007, LNCS 4484, 296–305, 2007.

[5] Cheng, Q., Wan, D. On the list and bounded distance decodability of Reed-Solomon codes. SIAM J. Comput. 37(1),195–209, 2007.

[6] Cheng, Q., Wan, D. Complexity of decoding positive-rate Reed-Solomon codes. IEEE Trans. Inform. Theory Vol. 56(10),5217–5222, 2010.

[7] Guruswami, V., Sudan, M. Improved decoding of Reed-Solomon and algebraic-geometry codes. IEEE Trans. Inform. Theory, Vol. 45(6) ,1757-1767, 1995.

[8] Guruswami, V., Vardy, A. A Maximum-likelihood decoding of Reed-Solomon codes is NP-Hard. IEEE Trans. Inform. Theory, Vol. 51(7), 2249–2256, 2005.

[9] R.Lidl and H.Niederreiter. Introduction to Finite Field and Their Applications, Cambridge University Press. 1986.

[10] Li, Y. J., Wan, D.: On error distance of Reed-Solomon codes. Science in China Mathematics Vol. 51(11), 1982–1988, 2008.

[11] Liao, Q. On Reed-Solomon Codes. Chinese Annals of Mathematics Series B 32B(1), 89–98, 2011.

[12] Sudan, M. Decoding of Reed-Solomon codes beyond the error-correction bound. J. Complexity 13, 180–193, 2007.

[13] Wu, R. and Hong, S. On deep holes of generalized Reed-Solomon codes. arXiv:1205.7016. 2012.

[14] Zhang, J., Fu F. W., Liao Q. Y. Deep holes of generalized Reed-Solomon codes(in Chinese). Sci Sin Maths, 43:727-740, 2013.

[15] Zhu, G., Wan, D., Computing Error Distance of Reed-Solomon Codes. TAMC 2012, LNCS 7287, pp. 214–224, 2012.